FDA Submission

Your Name: Dhawal Wazalwar

Name of your Device: HippoQuant – AI-powered hippocampal volume quantification system

Algorithm Description

1. General Information

Intended Use Statement:
HippoQuant is a software-only medical device intended to assist radiologists and clinicians by automatically segmenting and quantifying the volumes of the right anterior and posterior hippocampus in brain MRI scans. The system supports clinical decision-making by providing consistent volumetric measurements that can be reviewed within a standard PACS environment. The software processes each incoming study automatically, and all outputs are intended for interpretation by qualified healthcare professionals.

Indications for Use:
This software algorithm is intended to assist healthcare professionals by providing automated right hippocampus segmentation and volume estimation from T2-weighted brain MRI scans. The tool processes pre-cropped 3D brain MRI images focused on the hippocampal region. The outputs are intended to assist, not replace, clinical decision-making, and must be verified by a radiologist.

Device Limitations:

- Performance Variability Across Conditions

  The algorithm's diagnostic performance may vary depending on the specific condition. For instance:

  - **Input Scope:** Scope for this software is limited to pre-cropped 3D volumes around the hippocampus (HippoCrop). Therefore in case full brain T2 MRI scan is given as an input, results might differ.
  - **Laterality:** The training data does not include left hippocampus samples, which may lead to both false positives and false negatives when evaluating performance on the left hippocampal region.
  - **Scanner Variability:** Performance may vary with different MRI scanner vendors or protocols.

  These variations underscore the importance of clinical context and the necessity for radiologist oversight in interpreting algorithm outputs.

- Hardware Requirements

  The algorithm is optimized for GPU-accelerated environments to ensure timely processing. While it can operate on CPU-only systems, users may experience longer inference times, which could affect workflow efficiency in high-throughput settings.

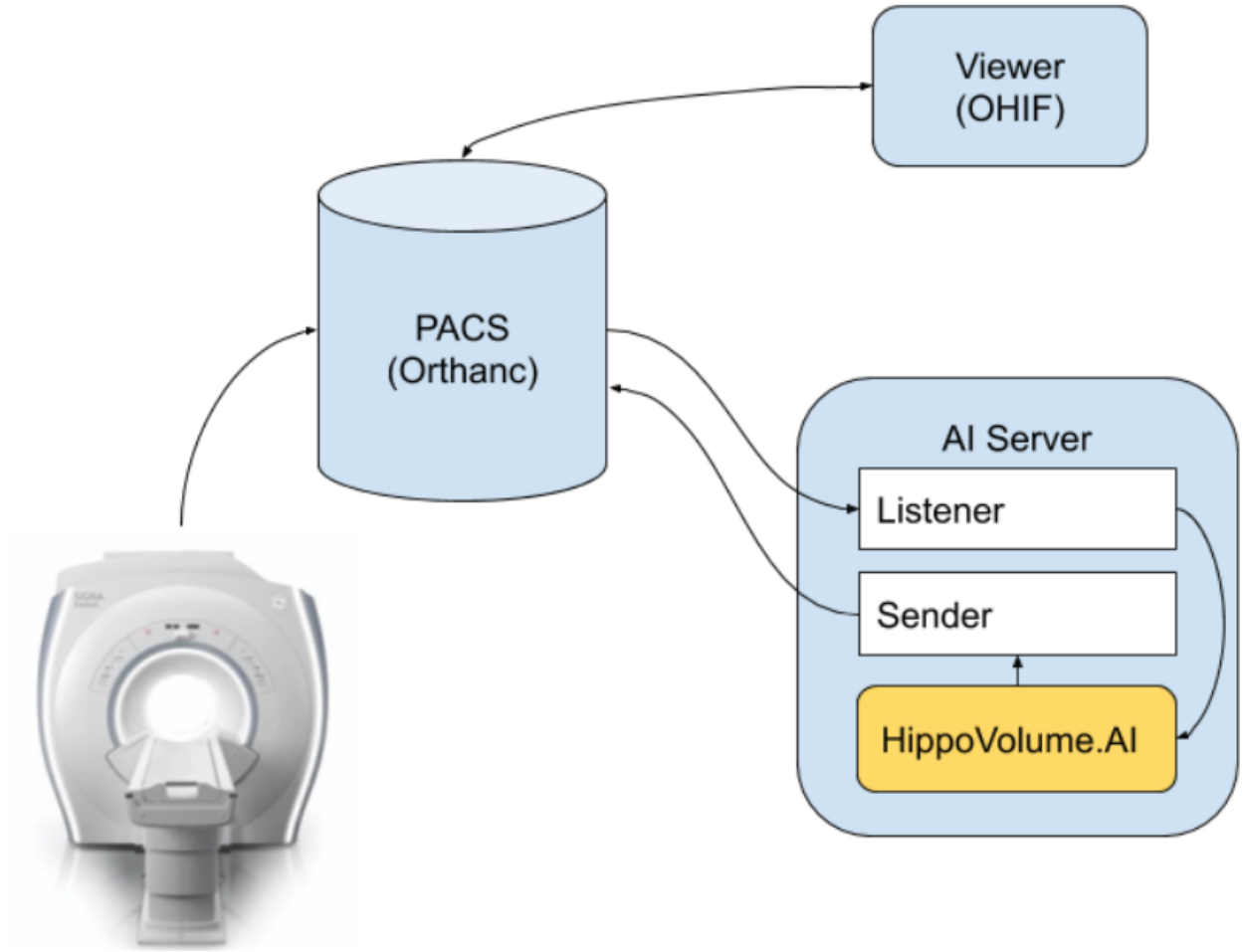
Clinical Impact of Performance:
Timely and automated measurement of hippocampal volume can assist in early diagnosis and monitoring of

neurological conditions such as Alzheimer's disease. Automated segmentation reduces inter-observer variability and saves clinician time. The clinical workflow benefit includes:

- Seamless integration into PACS
- Auto-reporting of hippocampal volume on every new relevant scan
- Flagging volumes below clinical thresholds for further review

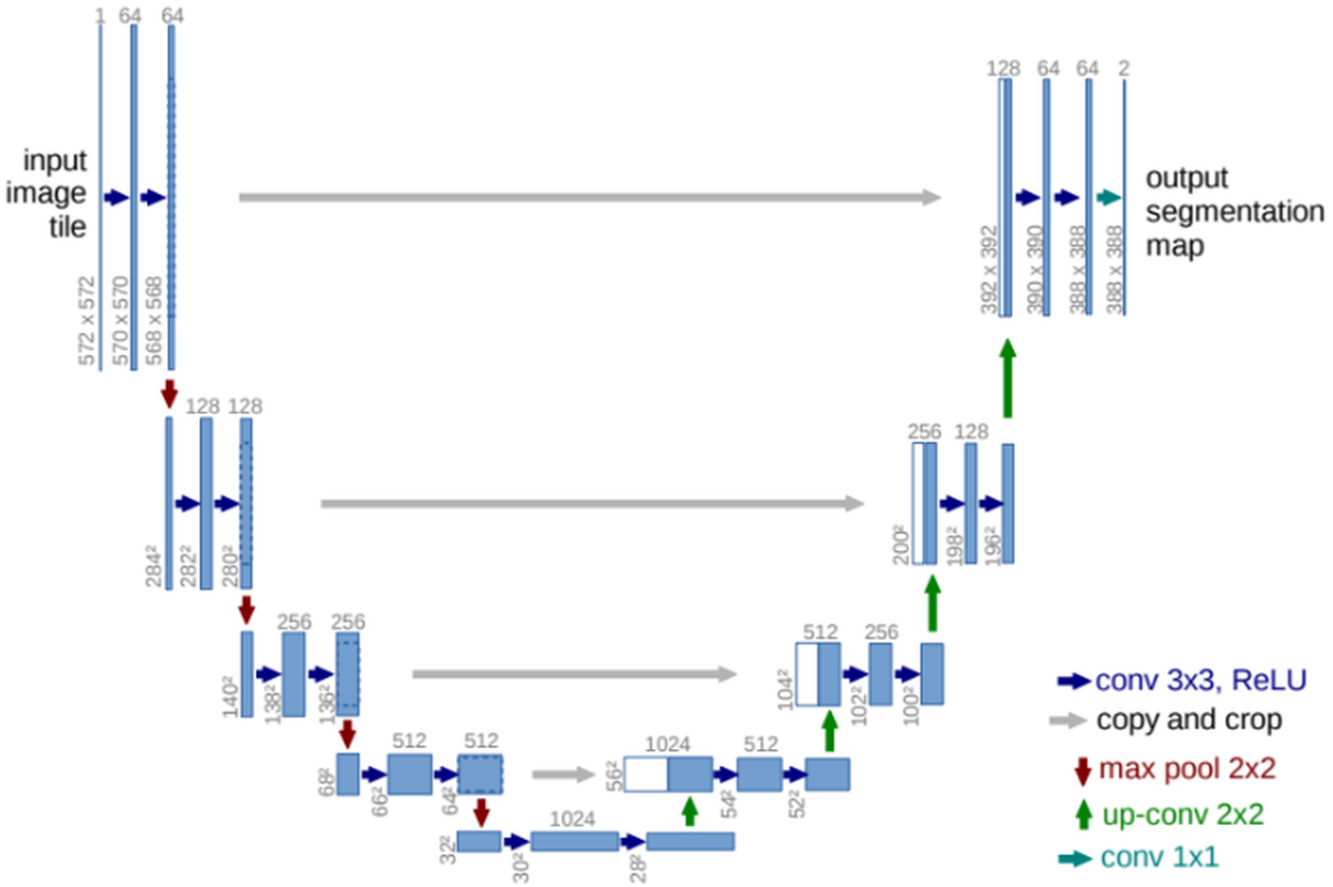
2. Algorithm Details and Clinical Integration

Clinical Flowchart:



Model Architecture:

- U-Net for 3D segmentation
- Input: 3D NIfTI volume (T2 MRI)
- Output: Binary mask for right hippocampus
- Loss: Dice Loss + Binary Cross-Entropy
- Framework: PyTorch



Inference Pipeline:

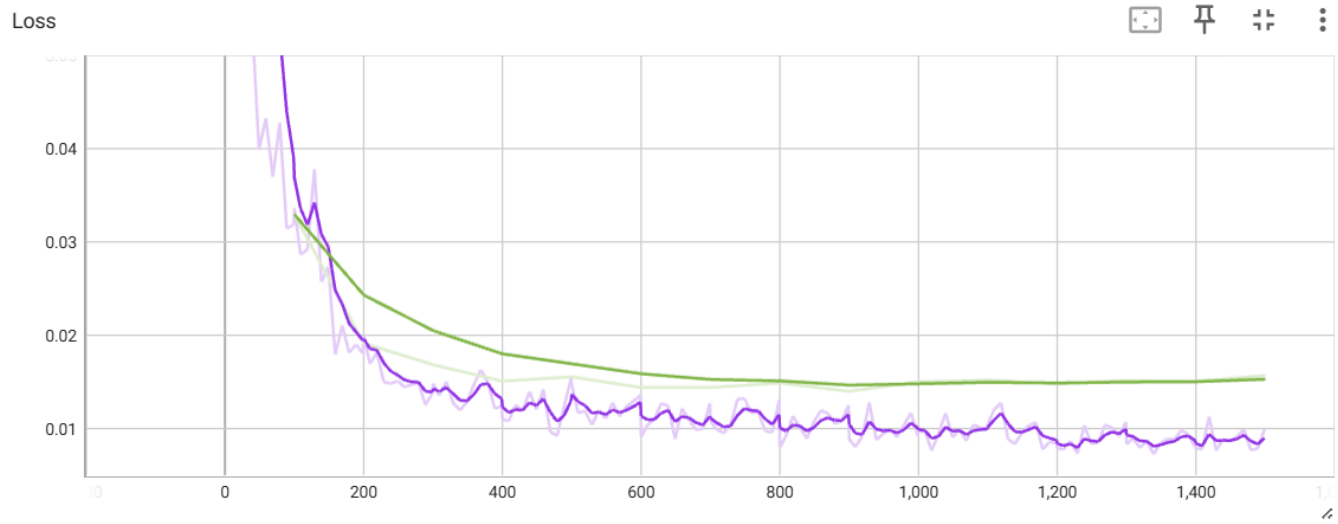
- Listen for new MRI study in PACS
- Identify series type (T2-weighted), confirm series description is "HippoCrop" and Modality is "MR"
- Run Inference and compute hippocampal volumes: Total, Posterior and Anterior
- Push result as structured report

Algorithm Training

Parameters:

- **Batch size:** 64
- **Optimizer:** Adam, Learning rate = 2e-4
- **Epochs:** 15
- **Patch Size:** 64

Training Visualization:



Run	Smoothed	Value	Step	Relative
May04_16-00-35_DW-Laptop_train	0.0091	0.0101	1,499	12.23 min
May04_16-00-35_DW-Laptop_val	0.0154	0.0158	1,500	11.51 min

Performance Evaluation on test data:

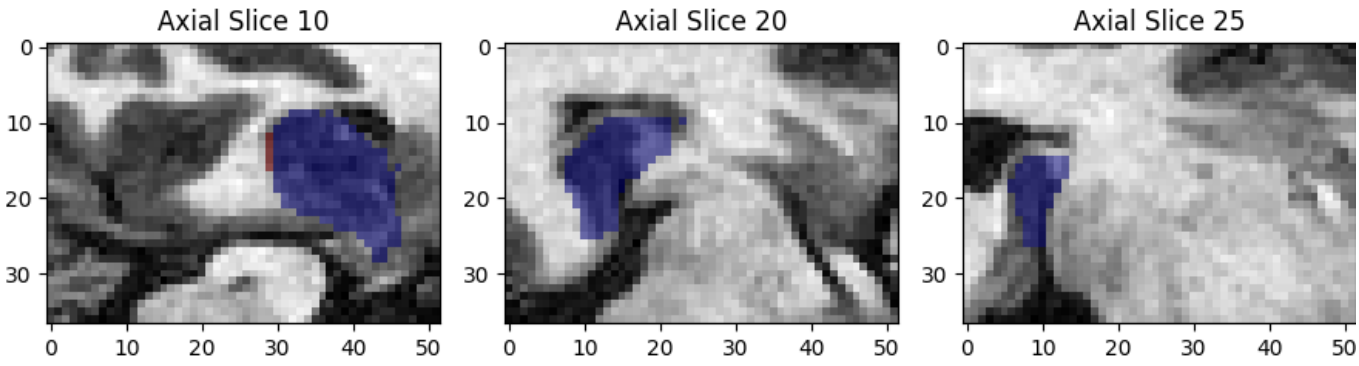
- **Mean Dice:** 0.9
- **Mean Jaccard:** 0.82

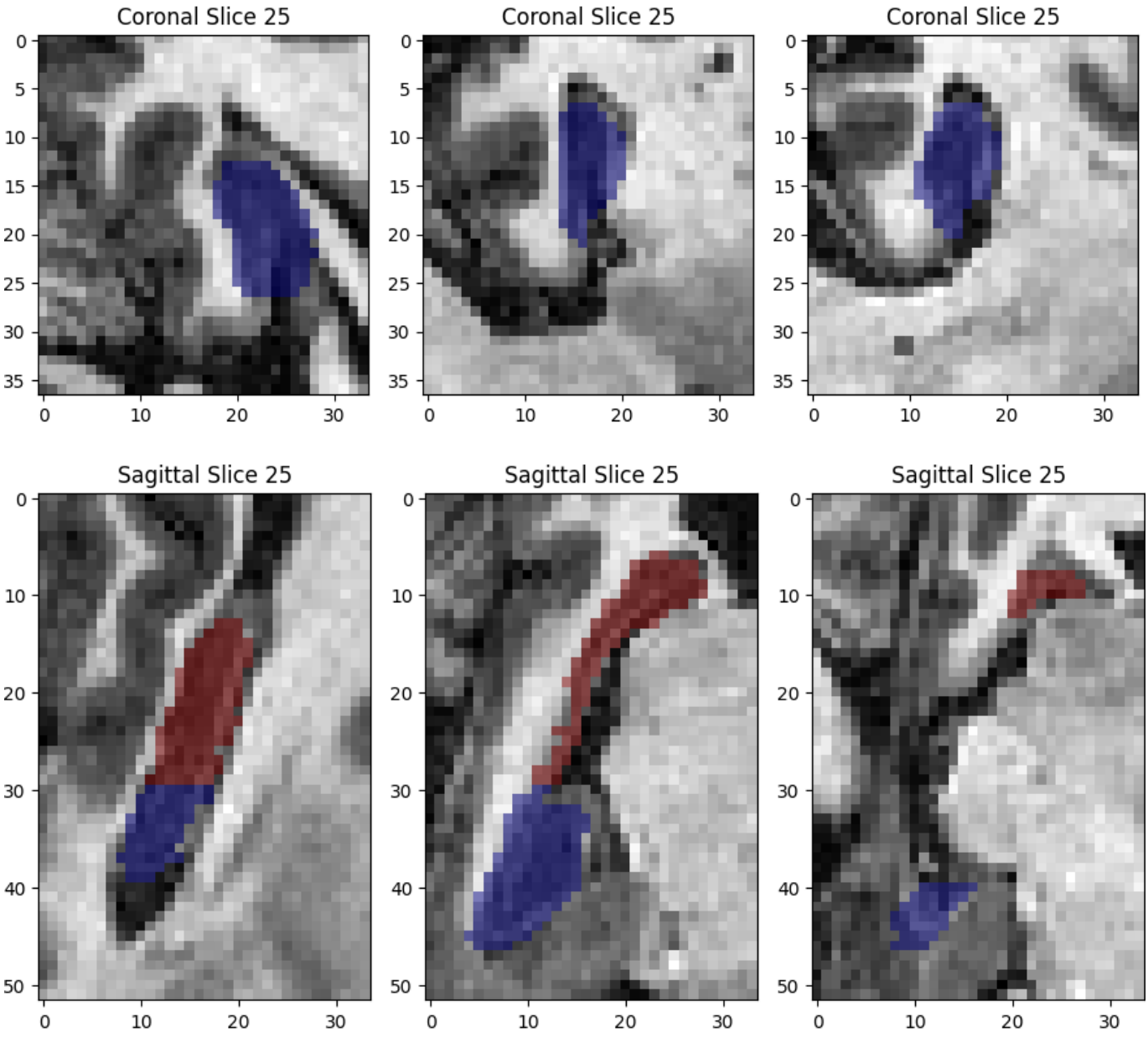
Data Description

Dataset used for this project is Medical Decathlon – Hippocampus dataset. This dataset is stored as a collection of 260 NIFTI files, with one file per volume, and one file per corresponding segmentation mask. The original images here are T2 MRI scans of the full brain. In this dataset we are using cropped volumes where only the region around the hippocampus has been cut out.

An 70:20:10 train-validation-test split strategy was used, with 70% of the data for training, 20% for validation and remaining 10% for test.

Below we can see some visualization of few slices for all 3 view: Axial, Sagittal and Coronal





Ground Truth

In this project, expert radiologists provided the annotations. They initially used the HippoCrop tool to extract a rectangular region from each brain scan series, performed the annotations on these cropped images, and subsequently converted them to NIFTI format. The resulting segmentation masks contain three distinct labels.

- **Background : 0**
- **Hippocampal Anterior : 1**
- **Hippocampal Posterior : 2**

FDA Validation Plan

Patient Population Description for FDA Validation Dataset:

- Patients aged between 18 and 100 years.
- The dataset should include a representative sample from various age groups, particularly focusing on those with conditions affecting hippocampal volume, such as Alzheimer's disease, but also including

healthy controls. Ideally, the dataset should have a balanced representation of abnormal and normal hippocampal volumes to help the model generalize across varied cases.
- The dataset will include T2-weighted 3D brain MRI scans, specifically pre-cropped around the hippocampal region (HippoCrop).
- The dataset should include scans from both male and female patients, ensuring gender balance.

Ground Truth Acquisition Methodology:
For reliable segmentation and volume quantification, the ground truth should be established using a multi-rater system:

- Segmentation of the hippocampus should be carried out by at least three expert radiologists or neurologists experienced with MRI scans of the brain.
- Each expert will manually segment the right anterior and posterior hippocampus, and their results will be cross-validated through a consensus-based approach or a majority voting system. If needed, weighted voting may be applied based on the radiologists' expertise.
- To ensure high-quality ground truth, radiologists should follow a standardized protocol to minimize variations due to manual segmentation discrepancies.

Algorithm Performance Standard:
The performance evaluation of HippoQuant should be conducted with a focus on two primary criteria:

- Dice Similarity Coefficient (DSC) for segmentation accuracy. A threshold DSC of 0.90 or higher should be expected between the algorithm's output and the radiologists’ consensus ground truth.
- Volume estimation accuracy, compared against manual measurements of hippocampal volumes provided by radiologists. The algorithm should achieve an absolute volume error of less than 10% compared to the manual estimates.